

**UNITED STATES DISTRICT COURT
FOR THE SOUTHERN DISTRICT OF NEW YORK**

**NATIONAL ASSOCIATION FOR THE
ADVANCEMENT OF COLORED PEOPLE,
SPRING VALLEY BRANCH, *et al.*,**

Plaintiffs,

v.

**EAST RAMAPO CENTRAL SCHOOL
DISTRICT, *et al.*,**

Defendants.

ECF CASE

Case No. 7:17-cv-08943

**DISTRICT JUDGE
CATHY SEIBEL**

**MAGISTRATE JUDGE
JUDITH C. MCCARTHY**

AFFIDAVIT OF DIRECT EXAMINATION OF DR. JOHN ALFORD

I. Background & Qualifications

1. My name is Dr. John Alford. I have been retained by counsel for the East Ramapo Central School District (the “District”) as an expert to evaluate the statistical analysis and conclusions of Dr. Steven P. Cole.

2. I am a tenured professor of political science at Rice University. I have worked as a professor for over 30 years. In that time, I have taught courses on redistricting, elections, political representation, voting behavior, and statistical methods at both the undergraduate and graduate level.

3. I have considerable experience performing and evaluating statistical analyses in the context of elections and voting rights litigation.

4. I have published over 40 articles about voting, elections, and political preferences, and have served as an article reviewer on these topics for over two dozen journals.

5. Over the last thirty years, I have worked with numerous local governments on districting plans and Voting Rights Act issues. I have previously provided expert reports and/or testified as an expert witness in over fifteen cases.

6. The details of my academic background, including all of my publications and work as an expert are included in my curriculum vitae, which is attached to this affidavit as Exhibit 1.

II. Scope of Review

7. I was retained to evaluate the analysis and conclusions of Dr. Steven P. Cole, who prepared a preliminary expert report and a rebuttal expert report in support of Plaintiffs' Motion for Preliminary Injunction in this case.

8. I reviewed both of Dr. Cole's reports as well as the sources to which he cites. I have also reviewed the transcript of Dr. Cole's first deposition, and the additional materials Dr. Cole produced to the District's counsel following that deposition.

9. In his reports, Dr. Cole reaches three primary conclusions: (1) Black voters in East Ramapo are politically cohesive, (2) there is a politically cohesive coalition of Black and Latino voters in East Ramapo, and (3) voting in East Ramapo is racially polarized.¹ In reaching those conclusions, Dr. Cole purportedly relied on four methods of analysis: King Ecological Inference (EI), a "correlation analysis," Homogeneous Precinct Analysis, and "Supplemental Research."²

10. In evaluating Dr. Cole's reports, I considered the reliability of each of the four methods. For each method, I sought to determine (1) whether the method could reliably be

¹ Preliminary Expert Report of Steven P. Cole, Ph.D. (Exhibit 2, Cole Report) at 2; Cole Deposition Transcript (Exhibit 3, Cole Tr.) at 45-46.

² Cole Report at 9-12; Rebuttal Expert Report of Steven P. Cole, Ph.D. (Exhibit 4, Cole Rebuttal Report) at 6.

used to evaluate voting patterns in the District, (2) whether Dr. Cole applied the method correctly, and (3) whether Dr. Cole's analysis of each method was reasonable. I also considered the analysis Dr. Cole used to reach his final conclusions.

11. I reach two primary conclusions regarding Dr. Cole's analysis. First, I conclude that Dr. Cole did not correctly apply his statistical methods, resulting in estimates and data that are not reliable.

12. Second, I conclude that Dr. Cole's results do not support his conclusions. More specifically, I conclude that:

- a. The results of Dr. Cole's EI analysis fail to show that Black voters vote cohesively;
- b. The results of Dr. Cole's EI analysis and Homogenous Precinct Analysis show that White voters vote for minority candidates at least as often as they vote for White candidates, and without variation in the degree of support according to the race of the candidate;
- c. Dr. Cole's correlation analysis does not support his finding of racially polarized voting;
- d. Dr. Cole's EI analysis of the exogenous 2012 U.S. presidential election is incongruent with his EI analysis of the school board elections because the 2012 presidential election was a partisan election with party nominated candidates, and had over seventy polling sites and much higher turnout;
- e. Assuming Dr. Cole's EI analysis of the 2012 presidential election is reliable, it shows that minority voters have the ability to elect their candidate of choice notwithstanding white bloc voting against the minority candidate of choice because President Obama won.

III. Ecological Inference

A. Background

13. The term “ecological inference” refers to the process of drawing inferences about individual behavior from aggregate data.³ In a general sense, ecological inference analyses are difficult to conduct because individual behavior often deviates from aggregate patterns.⁴

14. Until the mid-1990s, attempts to perform ecological inference analyses proved intractable. In the absence of individual-level data, it was difficult, if not impossible to make individual-level conclusions based on aggregate-level data.⁵

15. In 1997, Dr. Gary King introduced a promising technique that can provide accurate estimates of individual-level data. The technique was initially described in a book titled “A Solution to the Ecological Inference Problem: Reconstructing Individual Behavior from Aggregate Data.” I refer to this technique as “King’s EI,” or simply “EI.”

16. Over the years, there have been three main advancements associated with King’s EI.

17. The first advancement took place in the late ’90s when King initially introduced EI. In the years that followed, the political science and statistics communities identified a major problem with King’s method. At that time, EI could only be used to estimate voter support when there were two racial groups (e.g., White and Black) and two candidates. If there were more than two racial groups or candidates, then one would have to run an independent EI analysis for each race of interest and for each candidate of interest.⁶ King referred to this as an “iterative” approach to “RxC” estimation.

³ Gary King, A Solution to the Ecological Inference Problem at xv (1997).

⁴ *Id.*

⁵ *Id.* at xvi.

⁶ In practice, this involves simulating a two-race analysis by comparing the racial group of interest against a “dummy” group comprising the combination of all the other races. So instead of comparing the Black population

18. Statisticians and social scientists (including King) quickly realized that the iterative approach to RxC estimation was methodologically flawed and could not be relied upon to generate reliable estimates for elections with more than two races of interest or more than two candidates.⁷

19. The second advancement associated with King's EI method took place in 2001. After acknowledging the shortcomings of the iterative RxC approach, King published a more advanced theoretical approach to RxC estimation. At the time, the theoretical approach was viewed as computationally impractical—it could take more than a week to run a single model on the computers available at that time. Because of these computational difficulties, King included an alternate method that could approximate the results of the theoretical approach in a timely fashion.⁸ Because the alternate method could accommodate elections with several races and/or candidates, it quickly became the standard approach utilized by experts in voting rights cases.

20. Finally, in 2007 Olivia Lau, Ryan T. Moore, and Michael Kellerman introduced a software module called “eiPack.” EiPack took advantage of advancements in computing technology and allowed political scientists to apply King's advanced technique directly, without approximation.⁹ I refer to this as the “third-generation” of EI. Others refer to it as the Bayesian Method or the RxC Method.

against the White population (as one would do if there were *actually* only two races of interest), one would compare the Black population against the *combination* of the White and Latino populations. Then, one would repeat this process for each other race and candidate of interest. So in an election with three candidates and three races of interest, one would have to run the first-generation method 9 times (i.e. once for each race-candidate pair).

⁷ See Karen Ferree, *Iterative Approaches to $R \times C$ Ecological Inference Problems: Where They Can Go Wrong and One Quick Fix*, 12 Political Analysis 143 (2004) (Exhibit 5).

⁸ See Ori Rosen, Wenxin Jiang, Gary King, & Martin A. Tanner, *Bayesian and Frequentist Inference for Ecological Inference: The $R \times C$ Case*, 55 Statistica Neerlandica 134 (2001) (Exhibit 6).

⁹ See Olivia Lau, Ryan T. Moore, and Michael Kellermann, *eiPack: Ecological Inference and Higher-Dimension Data Management*, 7 R News 43 (2007) (Exhibit 7).

21. In certain settings, all three of the EI methods (and potentially other estimation methods as well) can produce reliable results. Specifically, in elections with two candidates of interest and two races of interest, and with a large number of voting precincts, the various EI methods produce similar or identical results.

22. In fact, in nearly every case where I have served as an expert, my EI estimates have been virtually identical to estimates generated by the opposing expert.

B. Dr. Cole's Analysis

23. Dr. Cole used EI analysis to estimate the proportion of voters, by race, who supported each school board candidate in each of the contested school board elections from 2013-2017. I evaluated Dr. Cole's EI analysis. I have identified several flaws in his methods and I disagree with his conclusions.

24. First, Dr. Cole did not use a good data set, because he did not start with the number of *actual voters* from each racial group when conducting his EI analysis.¹⁰ Instead, he started with estimates of the citizen voting age population (CVAP) for each precinct, as generated by Plaintiffs' expert demographer William S. Cooper.¹¹ The problem with that approach is that it assumes, without justification, that racial groups vote in proportion to their size—i.e., that if Blacks comprise 30% of the voting age population, then 30% of the votes cast in the election will be cast by black voters. In fact, studies have disproven that assumption. Typically, Black and Latino populations have significantly lower turnout than White voters.¹²

¹⁰ Cole Report at 7.

¹¹ See *id.*

¹² See, e.g., Kimball Brace & Bernard N. Grofman et al., *Minority Voting Equality: The 65 Percent Rule in Theory and Practice*, 10 Law & Policy 43-44 (1988) (Exhibit 8) (recognizing the "frequent observation that minorities customarily vote at lower rates than do whites").

Dr. Cole did not acknowledge this assumption, nor did he take any action to modify Dr. Cooper's CVAP estimates to account for likely differences in voter turnout by race.

25. The problems associated with using CVAP as a proxy for turnout are well recognized in the literature. The problem, as well as potential solutions, are described in considerable detail in Bernard Grofman & Matt A. Barreto's *A Reply to Zax's (2002) Critique of Grofman and Migalski*.¹³ Grofman and Barreto explain: "[D]ata on the racial composition of the electorate at the level of [precincts] are also rarely available. Thus, expert witnesses will normally have to use minority and nonminority shares of voting age population (VAP) or of registration as proxies for the minority and nonminority composition of the actual voting electorate. But when they do so, they are obviously inputting error in the nature of the independent variable. Furthermore, it can be shown that the error in estimates so generated is of a nonlinear form."¹⁴ Gary King has also recognized this problem.¹⁵

26. In conducting his analysis, Dr. Cole should have accounted for this difficulty. The most widely used technique for dealing with this issue is some form of double equation regression or double equation EI as discussed and advocated in detail in the Grofman and Barreto article that Dr. Cole cites in his rebuttal.¹⁶ Dr. Cole could also have used a more modern true RxC approach for his EI estimation (that is, third-generation EI). In a true RxC approach, like the one I used in this case, turnout differences are captured in the model by

¹³ Bernard Grofman & Matt A. Barreto, *A Reply to Zax's (2002) Critique of Grofman and Migalski*, 37 Sociological Methods & Research 599 (2009) (Exhibit 9). Dr. Cole cites this article in his rebuttal report to bolster the credibility of his analysis. In fact, the article shows that his analysis incorporated unreasonable assumptions.

¹⁴ *Id.* at 600.

¹⁵ See Gary King, *A Solution to the Ecological Inference Problem* 94 (1997).

¹⁶ See *A Reply to Zax*, *supra* note 13 (Exhibit 9), at 602; see also King, *A Solution to the Ecological Inference Problem* 94 (1997) ("However, although data on [the proportion of votes cast by each race] are available in a few data sets, these data are quite rare in real voting and most other applications. Thus, almost any practical use of aggregate data in race and voting studies to make inferences about individuals should include the insights from the double regression procedure.").

including non-vote, along with the various candidates, in the possible choices a voter might make. Alternatively, Dr. Cole could have estimated voter turnout by race using surnames and voter sign-in records, rather than CVAP estimates.¹⁷ Any of these methods would help account for the flaw in CVAP data. Dr. Cole did not perform any of them.

27. In the absence of corrective steps or, at the very least, an explanation as to why no corrective action was needed, Dr. Cole's faulty assumption fatally undermines his analysis, as it calls into question the inputs he used to conduct his analysis. Moreover, given the low number of precincts, any shortcoming in the data could easily have an outsized effect on the resulting estimate, since each data point is more significant. Thus, even a small error can make a big difference in the end result.

28. The second problem with Dr. Cole's EI analysis is that he used an outdated and inappropriate version of EI.

29. As explained above, there are three generations of EI analysis. Dr. Cole used the first generation of EI.¹⁸ While that method does not *always* produce unreliable results, it is widely recognized that the method should not be used to generate estimates in elections involving more than two candidates of interest and/or more than two races of interest.

a. King acknowledged this shortcoming, and responded by expanding his method. See Gary King, Ori Rosen, & Martin A. Tanner, *Binomial-Beta Hierarchical Models for Ecological Inference*, 28 *Sociological Methods & Research* 61, 61-62 (1999) (Exhibit 11) (concluding that the updated-EI method "can reveal some features of the data that King's model does not"); Rosen & King et al., *supra* note 8 (Exhibit 6), at 135-36 (2001).

¹⁷ See Kosuke Imai & Kabir Khama, *Improving Ecological Inference by Predicting Individual Ethnicity from Voter Registration Records*, *Political Analysis*, 24 *Political Analysis* 263 (2016) (Exhibit 10); see also *United States v. Village of Port Chester*, 704 F. Supp. 2d 411, 431-32 (S.D.N.Y. 2010) ("Experts for both parties used the Census Bureau List of Spanish Surnames to calculate the number of Hispanic voters in a particular area Neither party disputes that Spanish Surname Analysis is an accepted methodology").

¹⁸ See Cole Tr. at 85; Cole Rebuttal Report at 9.

b. Other statisticians have recognized the problems with the technique. *See* Ferree, *supra* note 7 (Exhibit 5).

30. The reason why first-generation EI should not be used in this setting is that first-generation EI only looks at one candidate-race pair at a time. As a result, the method fails to take advantage of the information we have regarding the voting patterns of other races or for other candidates.

31. It is worth noting that the problems associated with first-generation EI are only problematic in certain circumstances. First, the methodological problems described above occur in limited contexts in which first-generation EI is used to estimate voting patterns regarding more than two races of interest. Those problems do not arise in a two-candidate, two-race system. Second, even in a setting with more than two races of interest, first-generation EI can provide reliable estimates of voting patterns if there is not much variation in voting behavior across different races.¹⁹ Third, the problems with first-generation EI are significantly magnified by a lack of sufficient data. In a setting with lots of data, first-generation EI (and a variety of other methods) can produce accurate results.

32. In defending his use of first-generation EI, Dr. Cole argues that it is a standard practice for social scientists to use first-generation EI to estimate voter behavior.²⁰ Dr. Cole bases that conclusion on two sources: an expert report from *United States v. Village of Port Chester*, 704 F. Supp. 2d 411 (S.D.N.Y. 2010) and an article by Bernard Grofman and Matt A. Barreto titled “A Reply to Zax’s (2002) Critique of Grofman and Migalski (1988): Double-Equation Approaches to Ecological Inference When the Independent Variable Is Misspecified.” These sources do not validate Dr. Cole’s approach. The EI analysis in both *Port Chester* and the

¹⁹ *See* Ferree, *supra* note 7 (Exhibit 5), at 146-47 (explaining when and why problems arise when first-generation is used to estimate voting behavior for more than two races of interest).

²⁰ Cole Rebuttal Report at 6-7.

Grofman article took place in a setting where first-generation EI analysis is reliable—in both instances, EI was used to analyze elections with only two candidates and only two races of interest. Grofman & Barretto, *supra* note 13 (Exhibit 9), at 603; *Port Chester*, 704 F. Supp. 2d at 416 (explaining that the VRA claim alleged that the at-large system “denied the Hispanic population” an equal opportunity to participate in the political process).

33. The elections considered in both Grofman and *Port Chester* contained a far greater number of precincts than are present in the District. In *Port Chester*, EI was used to analyze an election with 16 precincts. *Port Chester*, 704 F. Supp. 2d at 420. In Grofman’s paper, EI was used to analyze elections in Los Angeles, which had 1,730 precincts. Grofman & Barreto, *supra* note 13 (Exhibit 9), at 613 n.6. The expert in *Port Chester* did not and could not use EI to analyze school board elections in the village, since there were too few precincts. See *Port Chester*, 704 F. Supp. 2d at 411.

34. Dr. Cole also defends his use of first-generation EI by arguing that first-generation EI produces similar results to third-generation EI. Because I used third-generation EI to conduct my analysis, I address that argument below.

35. The third problem with Dr. Cole’s EI analysis is that he did not produce sufficient materials that would allow one to replicate his results. Dr. Cole used a computer program called “EzI” to perform his analysis, and appears not to have retained any record of the various commands and settings that he provided as input into the EzI program. It is my understanding that EzI allows users to store and document the data and settings used to run a particular EI analysis. I have reviewed all of the materials Dr. Cole generated in association with his EI analysis. Those materials do not include any files or information disclosing the particular

settings he used to perform his analysis. Instead, his files contain only unlabeled output sheets that show the results of his analysis, but do not reveal the settings that produced the results.²¹

36. While I am able to evaluate Dr. Cole's general method of analysis based on his report and deposition testimony, the lack of proper documentation makes it impossible for me to replicate his results. I am also unable to confirm that Dr. Cole ran the program correctly or that he used reasonable settings in conducting his analysis.

37. The fourth problem with Dr. Cole's EI analysis is his inappropriate treatment of measures of reliability in both of his reports.

38. The standard practice among social scientists and statisticians is to accompany specific EI estimates (referred to as "point estimates") with associated confidence intervals.

39. Confidence intervals are used as a general indicator of how confident we should be with the estimated value of a given point estimate. Confidence intervals include two components: a range of values and a confidence level, expressed as a percent. The standard practice among social scientists and statisticians is to construct confidence intervals at a 95%-level. If a 95%-confidence interval is properly constructed, we can say that we are 95% confident that the interval contains the true value of the quantity of interest (e.g., the true proportion of Black voter support for a candidate). The ability to calculate a confidence interval is included in all current software implementations of EI that I have seen.

40. When interpreting EI results, the standard practice among social scientists and statisticians is to rely on confidence intervals in conjunction with point estimates. For

²¹ I have attached two of Dr. Cole's output pages to this affidavit as Exhibit 12. The first page shows Dr. Cole's estimate of Black voter support for Bernard Charles in the 2013 election, and the second page shows Dr. Cole's estimate of Black voter turnout in the 2013 election.

example, if an EI analysis estimates that a candidate received 55% of the Black vote, with a 95% confidence interval between 53% and 57%, one could conclude, with a 95% level of confidence, that the candidate received a majority of the Black vote, since 50% falls outside of the entire 95% confidence interval. However, if the confidence interval for the same 55% point estimate ranged between 25% and 85%, then one would *not* be able to conclude that the candidate received a majority of the Black vote, since multiple possible values at or below 50% lie within the 95% confidence interval.

41. Dr. Cole's preliminary expert report did not include *any* measure of reliability for his estimates. At his deposition, Dr. Cole testified that he did not calculate confidence intervals.²²

42. In my expert declaration (attached as Exhibit 13), I criticized Dr. Cole for failing to include standard errors or confidence intervals with his estimates. I concluded that the lack of standard errors or confidence intervals rendered his results unreliable.²³

43. In his rebuttal report, Dr. Cole included standard error measurements for his results and used his standard error estimates to generate confidence intervals.²⁴ I reviewed Dr. Cole's confidence intervals and conclude that they suffer from serious deficiencies.

- a. The confidence intervals and standard errors were generated using the same flawed methodology that produced the underlying estimates. As explained above, first-generation EI does not function properly when it is used to estimate voting behavior for more than two races of interest or in races involving more than two candidates. Because Dr. Cole's standard errors and confidence intervals were calculated using the same method, they suffer from the same deficiencies.
- b. Dr. Cole's confidence intervals do not make any substantive sense. His confidence intervals for his estimates of Latino voter support

²² Cole Tr. at 80, 138, 157.

²³ Alford Declaration at 9.

²⁴ Cole Rebuttal Report at 28-30.

for *every* election span less than one percent (e.g., he is 95% confident that between 99% and 100% of Latino voters voted for Jean Fields in the 2016 election). That result is so extreme that it lacks facial validity – the point estimate is too high, and the confidence interval is so narrow that it is not even an interval – it is just his point estimate. Besides that lack of facial validity, those confidence intervals don't make sense in terms of the data.

- (i) In the context of District elections, we have less information about Latino voters than we do about voters of any other race. Relative to the other races of interest, the Latino population has the lowest number of voters, the lowest proportion of voters, and the lowest amount of variation between precincts (both in terms of number of Latino citizens of voting age and in terms of the percentage of the voting age population). In fact, there was so little information about Latino voters, that Dr. Cole was unable to generate estimates for Latino voting patterns in two elections.²⁵ Despite these facts, Dr. Cole would have us believe that we should be *most* confident in his estimates relating to the group about which we have the *least* information. That is not how confidence intervals work. As explained above, all other things being equal, if you have less information, you should have correspondingly wider confidence intervals. Thus, the confidence intervals for White voters should be narrower than the confidence intervals for Latino voters. That they are not suggests there is a serious flaw in the way Dr. Cole constructed his confidence intervals.²⁶
- (ii) Just as troubling is the fact that *none* of Dr. Cole's confidence intervals includes the estimate of White support determined through homogenous precinct analysis (HPA). Given Dr. Cole's conclusion that HPA provides a reliable estimate of White voting patterns, the inconsistency between his EI confidence intervals and his HPA estimates strongly suggests a serious problem with either or both methods of analysis.

²⁵ See Cole Report at 39; Cole Tr. at 131-33.

²⁶ This problem also applies to Dr. Cole's confidence intervals for the Black population in the 2016 election. The confidence intervals for Black voting patterns in each 2016 election each span less than 3%. Four of the six estimates span less than 1%. These confidence intervals are much narrower than the intervals for White voting patterns despite the fact that we have much more data regarding White voters and White voting patterns than we do Black voters and Black voting patterns.

44. In evaluating Dr. Cole's preliminary expert report, I conducted my own EI analysis to see if I could replicate his results and to determine whether an EI analysis of District elections would be accompanied by wide confidence intervals.

45. I conducted my EI analysis using the third-generation version of EI. The method I used is fully described in Ori Rosen, Wenxin Jiang, Gary King, & Martin A. Tanner, *Bayesian and Frequentist Inference for Ecological Inference: the $R \times C$ Case*, 55 *Statistica Neerlandica* 134 (2001) (Exhibit 6). I ran my EI analysis in R using a software package/module called ei.MD.bayes, called from the widely used package of EI routines called eiPack.²⁷

46. I regularly use this method of EI analysis when consulting or testifying in voting rights cases. None of my EI analyses has ever been deemed unreliable or excluded by a court.

²⁷ See Lau et al., *supra* note 9 (Exhibit 7).

47. The results of my EI analysis are reproduced in the table below.

Candidate	Election Date	Race	White Vote %	Black Vote %	Latino Vote %
Mark Berkowitz*	5/16/17	W	86% (75, 95)	54% (12, 90)	73% (40, 95)
Alexandra K. Manigo	5/16/17	W	14% (5, 25)	46% (10, 88)	27% (5, 60)
Harry Grossman*	5/16/17	W	86% (75, 96)	53% (10, 89)	26% (5, 60)
Eric Goodwin	5/16/17	B	14% (4, 25)	47% (11, 90)	74% (40, 95)
Joel Freilich*	5/16/17	W	87% (77, 96)	59% (14, 92)	31% (7, 65)
Chevon Dos Reis	5/16/17	L	13% (4, 23)	41% (8, 86)	69% (35, 93)
Candidate	Election Date	Race	White Vote %	Black Vote %	Latino Vote %
Bernard L. Charles, Jr.*	5/17/16	B	85% (74, 95)	59% (13, 91)	34% (7, 69)
Kim A. Foskew	5/17/16	W	15% (5, 26)	41% (9, 87)	66% (31, 93)
Pierre Germain*	5/17/16	B	85% (74, 94)	58% (14, 92)	32% (6, 66)
Jean E. Fields	5/17/16	B	15% (6, 26)	42% (8, 86)	68% (34, 94)
Yehuda Weissmandl*	5/17/16	W	84% (73, 94)	56% (11, 92)	27% (5, 60)
Natashia E. Morales	5/17/16	L	16% (6, 27)	44% (8, 89)	73% (40, 95)
Sabrina Charles-Pierre	5/17/16	B	N/A	N/A	N/A
Candidate	Election Date	Race	White Vote %	Black Vote %	Latino Vote %
Jacob L. Lefkowitz*	5/19/15	W	82% (70, 93)	42% (6, 81)	17% (3, 43)
Sabrina Charles-Pierre	5/19/15	B	16% (6, 29)	53% (15, 89)	75% (45, 92)
Alan Keith Jones	5/19/15	B	1% (0, 3)	5% (1, 16)	8% (2, 20)
Yonah Rothman*	5/19/15	W	82% (69, 94)	48% (8, 86)	20% (3, 52)
Natasha Morales	5/19/15	L	18% (6, 31)	52% (14, 92)	80% (48, 97)
Juan Pablo Ramirez*	5/19/15	L	78% (65, 89)	47% (8, 84)	21% (4, 50)
Steve D. White	5/19/15	W	17% (6, 30)	49% (12, 87)	74% (43, 93)
Yisroel Eisenbach	5/19/15	W	6% (3, 8)	5% (0, 15)	5% (1, 14)
Candidate	Election Date	Race	White Vote %	Black Vote %	Latino Vote %
MaraLuz Corado*	5/21/13	L	84% (67, 96)	50% (10, 87)	23% (4, 54)
Margaret Tuck	5/21/13	B	17% (4, 33)	50% (13, 90)	77% (46, 96)
Pierre Germain*	5/21/13	B	80% (67, 92)	52% (11, 88)	26% (5, 59)
Eustache Clerveaux	5/21/13	B	20% (8, 33)	48% (12, 89)	74% (41, 95)
Bernard L. Charles, Jr.*	5/21/13	B	80% (66, 92)	52% (11, 89)	25% (5, 57)
Robert Forest	5/21/13	B	20% (8, 34)	48% (11, 89)	75% (43, 95)

48. The table contains six columns. The first column lists the name of each candidate who ran for a contested seat on the Board between 2013 and 2017. Consecutive rows are shaded to indicate which candidates ran against each other. For example, in 2013, MaraLuz Corado ran against Margaret Tuck, Pierre Germain ran against Eustache Clerveaux, and Bernard L. Charles, Jr. ran against Robert Forest. The winner of each election is indicated with a “*”.

The second column lists the date of each election. The third column indicates the race of the

candidate. The fourth column contains the point estimate and 95% confidence interval for the percent of white voters who voted for each candidate. The fifth column contains the point estimate and 95% confidence interval for the percent of Black voters who voted for each candidate. The sixth column contains the point estimate and 95% confidence interval for the percent of Latino voters who voted for each candidate.

49. The most important thing to recognize from my EI analysis is that each of the estimates relating to Black and Latino voters is accompanied by a very wide confidence interval.

50. Indeed, in the two-candidate races, *all* of the confidence intervals relating to Black voter support include fifty-percent. That means that, for each election, the EI results cannot tell us, with 95% confidence, which candidate received a majority of the Black votes.

51. What's more, in the two-candidate races, *all* of the confidence intervals relating to Black voter support span over seventy-five percent of all possible outcomes. For example, the most reliable estimate of Black voter support (that is, the estimates accompanied by the smallest confidence interval) indicate that Pierre Germain received somewhere between 11% and 88% of the Black vote in 2013. A conclusion that Germain received 88% of the Black vote would indicate that Black voters gave cohesive support to him, while a conclusion that Germain received 11% of the Black vote would indicate that Black voters gave cohesive support to his opponent. Most of the possible range between these endpoints would indicate, as would the point estimate of 52%, that Black voters did not give cohesive support to either candidate in this contest. In practical terms, that tells us nothing about the level of Black support, other than the fact that Mr. Germain received at least *some* Black support, but did not receive literally *all* of the Black votes. The same is true for every other candidate included in the EI analysis (with the

exception of Alan Keith Jones and Yisroel Eisenbach, who each participated in a three-candidate election and received a very small fraction of all votes—the EI analysis allows us to conclude, with 95% confidence, that both Mr. Jones and Mr. Eisenbach each received between ~0-15% of the Black vote, and a similar low percentage of support from Latino and White voters).

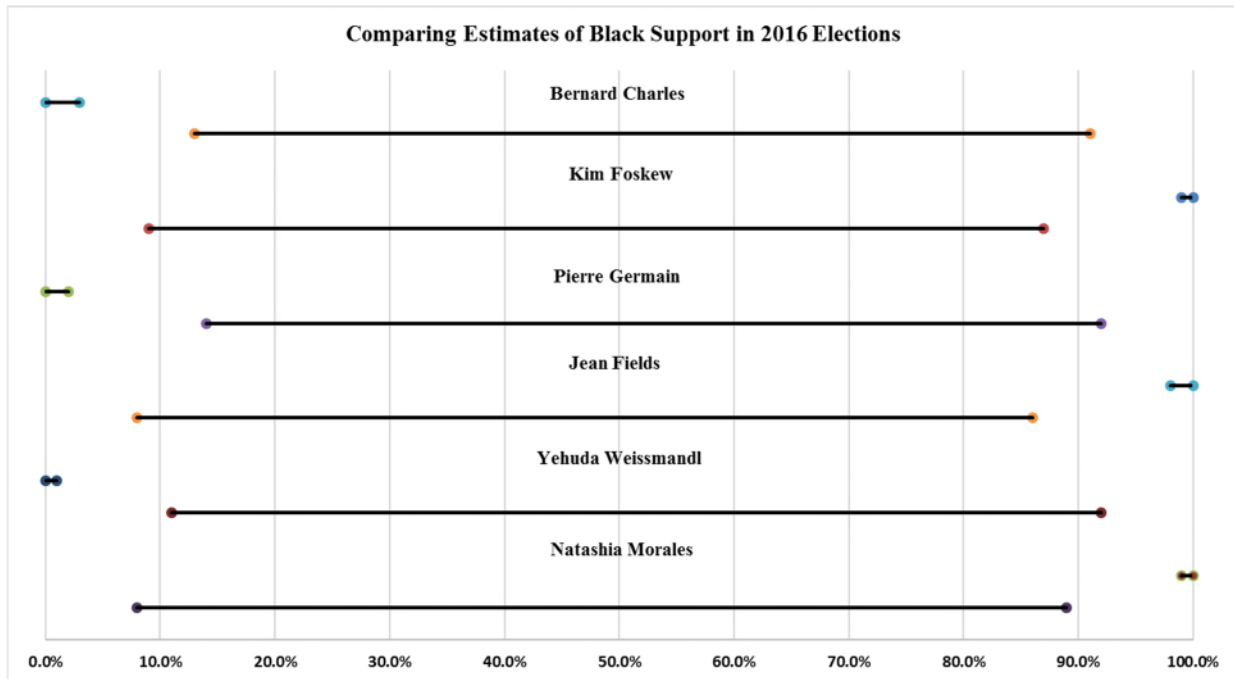
52. The confidence intervals are not much better when it comes to the point estimates of Latino voter support. In the two-candidate races, *all* of the confidence intervals relating to Latino voter support include fifty-percent. Once again, this means that the EI results cannot tell us, with 95% confidence, which candidate received a majority of the Latino votes.

53. In the two-candidate races, *all* of the confidence intervals relating to Latino voter support span over forty-five percent of all possible outcomes.

54. The confidence intervals produced in my analysis contrast with the confidence intervals contained in Dr. Cole’s rebuttal report. In my analysis, the confidence intervals for White voters are considerably narrower than the confidence intervals for Black and Latino voters. As explained above, this is what one would expect to see, given that we have significantly more information about white voters than we have about black or Latino voters.

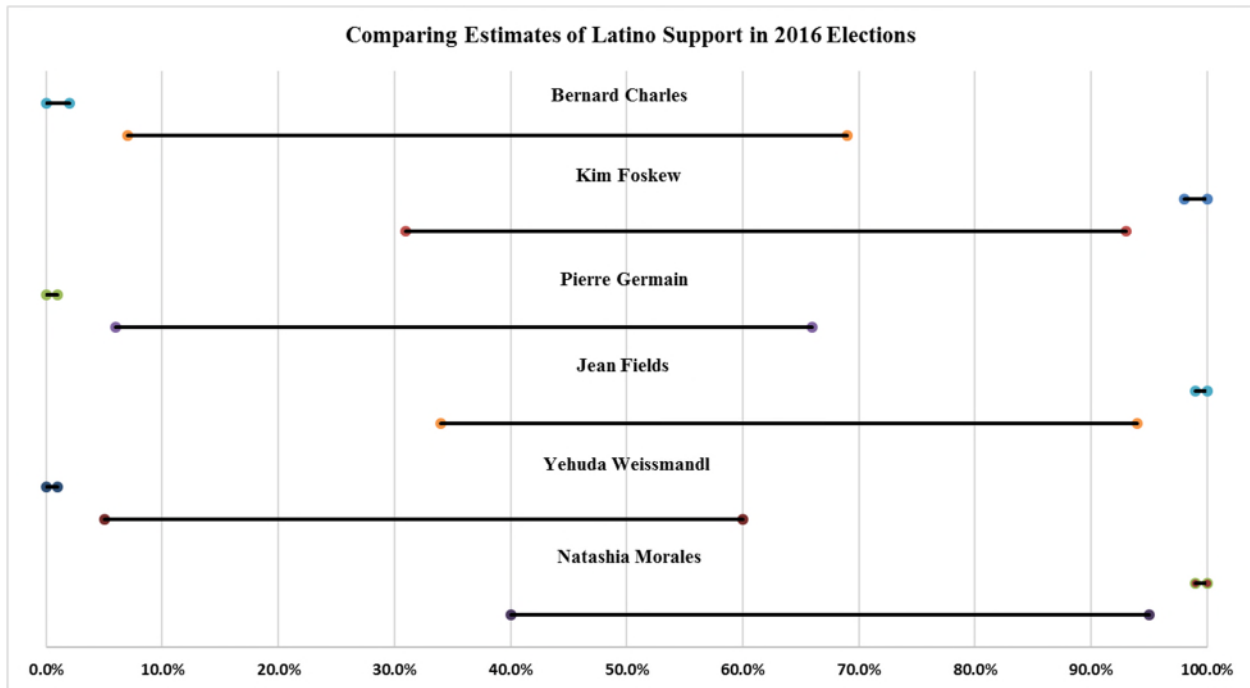
55. The differences between my confidence intervals and Dr. Cole's confidence intervals can be seen visually in the following graph, which shows my and Dr. Cole's 95% confidence intervals for Black voter support in the 2016 election. Our respective intervals for the other elections diverge in essentially the same way.

56. The top line in each set shows the confidence interval generated by Dr. Cole. The bottom line in each set represents my confidence interval.



57. Our confidence intervals for Latino candidates exhibit similar differences.

The following graph shows my and Dr. Cole's 95% confidence intervals for *Latino* voter support in the 2016 election. Once again, the top line in each set shows the confidence interval generated by Dr. Cole, while the bottom line in each set represents my confidence interval.



58. It is also useful to look at where my EI analysis yields results that are at least broadly consistent with Dr. Cole's EI analysis. In particular, both my EI analysis and Dr. Cole's EI analysis show that, in each election, White voters had a clear candidate of choice.²⁸ Moreover, our respective EI analyses agree on who that candidate of choice was in each election. While these combined observations do not allow us to draw conclusions regarding minority voting behavior, they allow us to conclude that a majority of White voters have voted for

²⁸ This is evident both from our respective point-estimates and from the fact that both of our analyses generated relatively narrow confidence intervals for White voting behavior. While our confidence intervals are still larger than ideal, they suggest that our analyses can be used to draw limited conclusions regarding White voter preferences. For example, my EI analysis indicates that in the 2017 election, we can say with 95% confidence that Mark Berkowitz received between 75% and 95% of the White vote. That means that we can assert, with confidence that Mr. Berkowitz won the White vote by a considerable margin. Dr. Cole's analysis supports the same conclusion, showing, at a 95% confidence level, that between 75% and 79% of White voters supported Mr. Berkowitz.

minority candidates in 2013, 2015, and 2016. They also show that a majority of White voters have sometimes supported Black or Latino candidates over White candidates, and that White voter support for candidates does not seem to vary according to the race of the candidate.

59. While I cannot offer a definitive explanation for the wide confidence intervals associated with my EI analysis of Black and Latino voters' preferences, I think it is likely that the EI analysis yielded wide confidence intervals because of the relatively low number of Black and Latino voters and because of the relatively low number of voting precincts in the District. In this sense, the wide confidence intervals are functioning here, as they should, as a diagnostic tool that indicates that EI does not provide much information about minority voting behavior in the District's school board elections.

60. Comparing my confidence intervals with Dr. Cole's EI point-estimates further demonstrates the wide disparity between our respective estimates of Black and Latino voter preferences. Despite the fact that the confidence intervals associated with my EI analysis are wide, the vast majority of Dr. Cole's point estimates of Black and Latino voting behavior nevertheless fall *outside* of the corresponding confidence intervals generated by my EI analysis.

61. With respect to two-candidate races, Dr. Cole generated 20 estimates of Black voter support (one estimate for each of two candidates in ten elections). 14 of those 20 estimates fall outside of the corresponding confidence interval generated by my analysis.

62. For example, my confidence interval associated with the 2016 race involving Bernard Charles indicates that somewhere between 13% and 91% of Black voters voted for Mr. Charles. However, Dr. Cole estimates that 1.4% of Black voters voted for Mr. Charles. Thus Dr. Cole's estimate falls outside of the confidence interval generated by my analysis.

63. With respect to two-candidate races, Dr. Cole generated 20 estimates of Latino voter support. *All* of Dr. Cole's estimates of Latino support fall outside of the corresponding confidence interval generated by my analysis. That is, *none* of Dr. Cole's estimates falls within the corresponding confidence interval generated by my analysis.

64. In his rebuttal report, Dr. Cole criticizes my observations that "it is my experience that EI analysis tends to produce unreliable estimates when there are fewer than 15 precincts" and that "the relatively low number of voters in each precinct ... would ... make it difficult to obtain reliable EI estimates."²⁹ Dr. Cole does not actually engage with my concern. My observations regarding the low number of precincts and voters is not, in and of itself a criticism of Dr. Cole's analysis. Instead, my observations, which are based on decades of experience using EI, provide a real-world, common sense explanation as to why EI likely would not tell us much about voting patterns in the District's school board elections. As can be seen from the confidence intervals, I was right to be concerned.

65. Instead of showing how EI can accommodate a low number of precincts, pointing to steps he took to account for the lack of data, or offering any other explanation for why his confidence intervals differed from mine, Dr. Cole simply identifies other school districts in New York that contain a low number of precincts and misrepresents my opinion. I have never stated, and it is not my opinion, that "voters in jurisdictions with a smaller number of polling places are less entitled to relief from vote dilution as voters in jurisdiction with a larger number of polling places."³⁰ No reasonable person could reach that conclusion from a fair reading of my declaration.

²⁹ See Alford Declaration (Exhibit 13) ¶¶ 15, 26.

³⁰ Cole Rebuttal Report at 14.

66. My opinion that Dr. Cole’s EI analysis produced unreliable results simply means that, in my view, Dr. Cole’s EI analysis produced unreliable results. My opinion that Dr. Cole’s EI analysis was flawed does not mean that it is impossible to determine voter behavior in the District. It is possible that a better-designed application of EI could generate reliable estimates, as could different statistical tools or methods of estimation (for example, one could estimate voting behavior by conducting an exit poll).

67. In his rebuttal report, Dr. Cole responds to my EI analysis by pointing to two articles that purportedly compare third-generation EI to first-generation EI and that purportedly conclude that the two methods produce “nearly identical results.”³¹ Even if that were true—and as discussed below there are problems with that assertion—it is a *non-sequitur*. The results of Dr. Cole’s first-generation EI analysis are demonstrably *not* “nearly identical” to the results generated by my third-generation EI analysis. Dr. Cole offers no explanation for that difference that rehabilitates his analysis.

- a. The two articles upon which Dr. Cole relies do not appear to actually use the RxC routine that I use in my analysis (ei.MD.bayes). Instead of comparing the first-generation EI method to the Bayesian method of ecological inference (i.e., the third-generation of EI), the articles compare the first-generation EI method to the Bayesian method of ecological *regression*—a technique that was developed in the early 1950s by Leo Goodman.³² See *eiCompare*, Collingwood et al., at 96 (“The RxC

³¹ Cole Rebuttal Report at 8.

³² The two articles upon which Dr. Cole relies were written by the same authors (Matt A. Barreto, Loren Collingwood, Kassra Oskooii, and Sergio Garcia-Rios) and appear to apply the same methods. The first article, titled *eiCompare: Comparing Ecological Inference Estimates Across EI and EI: R x C*, introduced a new R package that can be used to compare different methods of performing ecological inference (including, but not limited to, King’s various EI methods). Collingwood et al., *eiCompare*, 8 The R Journal 92 (2016) (Exhibit 14). It is easy to determine how the authors conducted their analysis because the article contains a copy of the script they used to run their comparison. *Id.* at 96. That is how I was able to determine that they used the wrong function call (ei.reg.bayes) in their analysis.

The second article, titled *Estimating Candidate Support: Comparing EI & EI-RxC Methods* (Exhibit 15), does not contain a copy of the script used to perform the analysis. However, the article considers the same elections that were considered in the first article, and contains an *identical* data output as the analysis in the first article—

builds off of code from the eiPack package [H]ere we use the ei.reg.bayes() function”).³³

- b. The data sets analyzed in the EI comparison articles were screened to make sure that they would be “amenable to ecological inference.”³⁴ Specifically, the authors created and analyzed tomography plots for “every single data set [they] considered” for inclusion in the article.³⁵ Methodologically, this decision makes sense—there is no point in comparing EI estimates if the available data means that EI would not generate accurate or reliable results. That is to say, there is no point in comparing EI estimates in a setting where EI would not be used in the first place.

Indeed, the authors recognize that “uninformative” data yield “standard errors that may be too large to be useful or simply incorrect.”³⁶ The fact that the cited articles pre-screened their data to ensure that EI would work means that their conclusions do not speak to whether first- and third-generation EI produce similar results in settings that are *not* amenable to EI analysis. Because the primary disagreement between me and Dr. Cole is whether the District elections are, in fact, amenable to EI analysis, the articles are of little value here (though they do support my earlier analysis challenging the accuracy of Dr. Cole’s confidence intervals). In other words, the articles show that first-generation EI can *sometimes* produce reliable results. That conclusion is well-established and beyond dispute. The question is whether this is such a setting, and the articles upon which Dr. Cole relies do not speak to that question.

- c. Along similar lines, it is worth noting that the authors of the cited articles had access to *far more* data than is available here. In fact, the smallest election considered in the articles had almost four times the number of precincts as are present here,³⁷ and several of

something I would only expect to see if the method of analysis were the same. Compare Table 6 in *eiCompare* with Table 8 in *Estimating Candidate Support*.

³³ R’s documentation confirms that the “ei.reg.bayes()” function “[e]stimate[s] an ecological regression using Bayesian normal regression.” *ei.reg.bayes*, RDocumentation, <https://www.rdocumentation.org/packages/eiPack/versions/0.1-7/topics/ei.reg.bayes>. The R documentation for that function call includes just one reference—to Leo Goodman’s 1953 article, titled “Ecological Regressions and the Behavior of Individuals.”

³⁴ Dr. Cole cites and relies upon a draft version of *Estimating Candidate Support*. The passage quoted above comes from page 11 of the final, updated version, dated March 27, 2017, which is located online at http://www.collingwoodresearch.com/uploads/8/3/6/0/8360930/smr_submission.pdf, and attached to this affidavit as Exhibit 15.

³⁵ *Id.* at 12.

³⁶ *Id.* at 11.

³⁷ See *Comparing Ecological Inference Estimates* (2017), at 31.

the elections they considered had *several hundred* times as many precincts. For many elections, they also knew exactly how many people of each race voted in each precinct.³⁸ This difference in data is significant, since first-generation EI is especially sensitive to a lack of data.

68. In his rebuttal report, Dr. Cole argues that “R x C [EI] has limitations” and that “the results do not always demonstrate face validity.”³⁹ I am not sure what Dr. Cole means by these statements. If he is arguing that RxC EI (i.e., third-generation EI) cannot always generate reliable estimates, then I am in full agreement—as my analysis in this case demonstrates, if there is a lack of sufficient data, then RxC estimates will be accompanied by wide confidence intervals. If, on the other hand, Dr. Cole means that there are circumstances in which first-generation EI produces more reliable estimates than third-generation EI, his argument is entirely without merit—Dr. Cole has not identified any instances in which first-generation EI has produced more reliable estimates than third-generation EI, nor has he offered a methodological explanation as to why an older, less advanced technique would produce better estimates and narrower confidence intervals.

69. In claiming that RxC results “do not always demonstrate face validity,” Dr. Cole is quoting the *eiCompare* article, which, as explained above, did not consider King’s RxC technique, but rather considered Goodman’s ecological regression. Moreover, that statement is inconsistent with the data actually contained in the article, which “demonstrate that ... EI and R x C produce remarkably consistent results” and show that, to the extent the estimates differ, the estimates generated by RxC are closer to true known values than those generated by first-generation EI.⁴⁰

³⁸ *Id.* at 9.

³⁹ Cole Rebuttal Report at 8.

⁴⁰ See *eiCompare*, *supra* note 34, at 99 & Table 6 (showing that the RxC estimates are closer to the exit poll estimates than the first-generation EI estimates).

70. Dr. Cole appears to mistakenly assume that I offered my EI analysis as an independent assessment of voting patterns in the District. That was not the purpose of my analysis, and I have offered no opinion on whether voting in the District is, or is not, racially polarized. The purpose of my EI analysis was to determine whether Dr. Cole's EI estimates were reliable. I was not asked to determine the actual proportion of voters who supported a candidate or to make an independent assessment about the presence of racially polarized voting.

71. Dr. Cole seems not to grasp this point, as is evident from his assertion that:

Dr. Alford's failure to generate any meaningful analysis regarding Black or Latino voting patterns in East Ramapo highlights the importance of following the generally accepted practice of using as many analytical methods as possible in evaluating racially-polarized voting. Dr. Alford effectively asserts that investigation into the existence of racially polarized voting should end if his one preferred method of EI yields standard errors too large or confidence intervals too wide. This is not well-accepted practice in this field, where analysts generally rely on as many measures as possible to triangulate voting patterns. I have followed that generally accepted practice.⁴¹

Dr. Cole's criticisms are *non-sequitur*. My EI analysis only sought to evaluate Dr. Cole's EI method. The other methods of analysis that Dr. Cole may (or may not) have used are irrelevant to that particular exercise. And, just to be perfectly clear, it is *not my opinion* that "that investigation into the existence of racially polarized voting should end if ... EI yields standard errors too large or confidence intervals too wide." Dr. Cole attacks a straw-man. There may be other methods for determining whether there is racially polarized voting here instead of EI. Dr. Cole did not use them, and I offer no opinion on the potential alternative methodologies that Dr. Cole might have used. My opinion is that Dr. Cole's EI analysis is flawed and unreliable because he used incomplete and inadequate data, he used an inappropriate, outdated, and largely

⁴¹ Cole Rebuttal Report at 16.

abandoned computer program, and his analysis has generated facially invalid and unreasonable estimates that no court should rely on. Nothing in Dr. Cole's rebuttal report changes my opinion.

IV. Goodman Single-Equation Ecological Regression

72. In his preliminary expert report, Dr. Cole listed "Goodman single-equation ecological regression" as one of the methods he used to analyze District elections. In my expert declaration, I challenged that assertion and explained that Dr. Cole did not *actually* perform a regression analysis, but instead simply generated a correlation coefficient between the racial composition of a district and the proportion of votes a candidate received in that district. In his rebuttal report, Dr. Cole confirms that he "did not ... use single regression estimates of racially polarized voting," but that he instead performed a correlation analysis.⁴²

V. Correlation Analysis

73. In his rebuttal report, Dr. Cole explains that he "used [a] correlation coefficient ("r") to measure the strength of the relationship between race of voters and vote outcome."⁴³

74. Dr. Cole calculated the correlation coefficient for each race of interest in each of the contested elections, and concludes that "[a] high r value demonstrates that race is tightly associated with voting behavior, which is an indication of racially polarized voting."⁴⁴

75. The largest problem with Dr. Cole's use of correlation coefficients is that none of his coefficients for Black voters in two-candidate elections was statistically significant at the .05 level, or even the .1 level. The lack of statistical significance means that we cannot trust that his correlation coefficients are accurate.

⁴² *Id.* at 18.

⁴³ *Id.*

⁴⁴ *Id.*

76. At his first deposition, Dr. Cole stated that he could rely on his correlation coefficients despite their lack of statistical significance because the correlation values were large.⁴⁵ That statement is directly at odds with widely recognized and adopted statistical principles. Statisticians and social scientists uniformly recognize that one cannot make up for a lack of statistical significance with a large effect size. Indeed, Dr. Cole has it backwards. The larger, or more anomalous a result, the *more* important it is to make sure that the observed measurement was not caused by error or random variance.

77. Dr. Cole cites and quotes Absoch et al.⁴⁶ in defense of the use of bivariate correlation analysis,⁴⁷ but Dr. Cole's use and interpretation of correlations in his report contradicts the explicit advice in the Absoch article. Dr. Cole relies on statistically non-significant correlation results that contradict his EI results. Absoch et al instead advise:

[The] first method (1) is to examine a series of bivariate correlations between proportions of voter preference for a particular candidate ... and the proportion of relevant Latino population within the same redistricting unit. *This is meant primarily to be an instructive device* – as the presence of high, *and statistically significant* correlations suggests, but may not be in isolation, conclusive evidence of racially polarized voting.⁴⁸

In keeping with that observation, every one of the correlation coefficients included in and relied upon in the Absoch article are significant at $p < .00$, well above the $p < .05$ standard. Dr. Cole argues for, and relies on, correlation coefficients that, in the case of his estimates for Black voters, consistently fail to meet the standard social science level of .05 (95% confidence).

⁴⁵ Cole Tr. at 193 (“Q. If I also recall correctly your correlation analysis for the 2013 election for the Black voters, you did not generate an estimate that was statistically significant, right? A. That’s correct, but the effect size of those correlations was of the magnitude that one can rely upon them.”)

⁴⁶ Yishaiya Absoch et al., *An Assessment of Racially Polarized Voting for and Against Latino Candidates in California*, in Voting Rights Act Reauthorization of 2006, at 107 (2007) (Exhibit 16).

⁴⁷ Cole Rebuttal Report at 1-2.

⁴⁸ *Absoch et al.* at 117 (emphasis added).

78. This is not to say that correlation analyses are useless. However, it would be inappropriate, as the advice of Bosch et al makes clear, to rely on a non-significant correlation analysis.

VI. Homogenous Precinct Analysis

79. In his preliminary expert report, Dr. Cole conducted a Homogeneous Precinct Analysis (HPA) using a “simple percentages” analysis.⁴⁹

80. A “homogenous precinct” is a precinct in which 90% of the population comes from the same race. According to Dr. Cole, in homogenous precincts, the proportion of the vote received by a candidate can function as an accurate estimate of the voting preferences of the homogenous group.⁵⁰

81. The District contains three homogenous districts—Lime Kiln, Kakiat, and Ramapo High School. Over 90% of the population in those districts is White.⁵¹

82. There are no homogeneous Latino or Black polling places in the District. As a result, HPA does not reveal anything about minority voting patterns in District elections. Thus, HPA cannot support the claim that minorities vote differently than whites. As was true for his use of bivariate correlation, Dr. Cole cites Absoch et al. to bolster his inclusion of homogenous precinct analysis, but what Absoch et al. lay out in their article is not what Dr. Cole provides here. Indeed, in explaining their analysis, Absoch et al. explain:

This method is probably the simplest method for examining polarized voting. We use redistricting units within District 3 that are either 90% non-Latino (or greater) or 90% Latino (or greater) and compare the two against each other. The ease with which this sort of comparison can be made, indeed without resorting to statistics of any kind, makes this a logical precursor to more sophisticated methods of analysis. A downside to this sort of analysis is the availability, or lack thereof, of units that are sufficiently homogenous to be compared. Also, depending on the political

⁴⁹ Cole Report at 9-10.

⁵⁰ *See id.*

⁵¹ *See id.*

jurisdiction in question, there may be some issue with assuming the patterns in more heterogeneous units will reflect what we see in homogenous ones.

Our analysis is a series of t-tests that statistically measure the difference between the two types of units in the level of support granted to Latino candidates and issues. A benefit of this sort of analysis is that we report the mean (or average) support within each type of homogenous unit, the difference, and associated standard errors, which allow for a determination of whether the levels of support are statistically discernible from each other.⁵²

83. Note that Absoch et al assess polarized voting by comparing the voting patterns in homogenous Latino precincts to the voting pattern in homogenous non-Latino precincts. This allows them to conclude that Latinos and non-Latino are supporting different candidates, and that this difference is statistically significant. Cole does neither of these. He is only able to examine homogeneous White precincts, and as a consequence has no estimation of which candidate non-Whites are supporting and no measure of statistical significance.

84. In his rebuttal report, Dr. Cole argues that his analysis of the 2012 presidential election supports his conclusions regarding the District's school board elections. I disagree. The 2012 election involved a larger electorate, an electorate that was more engaged, a completely different set of issues, partisan politics, and a different voting structure. Even if Dr. Cole's conclusions regarding those elections were correct, they would not impact his analysis of school board elections in either direction.

85. To the extent one thinks that the 2012 presidential election is relevant, the results from that election undermine Dr. Cole's conclusion that "candidates preferred by Black or a coalition of Black and Latino voters in the District are usually defeated by candidates supported by the White majority voting bloc."⁵³ Dr. Cole's analysis of the 2012 election shows that members of the minority community have the ability to elect their candidate of choice—his

⁵² *Absoch et al.* at 121-22.

⁵³ Cole Report at 1.

results show that President Obama was the minorities’ preferred candidate, that Mitt Romney was the whites’ candidate of choice, and that President Obama received a majority of votes in the election.⁵⁴

VII. “Supplemental Evidence” and Mixed-Methods Approach

A. Dr. Cole’s Newly Articulated “Mixed-Methods Approach” Lacks an Adequate Research Plan

86. In his preliminary and rebuttal expert reports, Dr. Cole states that he used “supplemental research” to understand the “factual context” of the Board elections, “and to assess the precision and accuracy of [his] quantitative analysis.”⁵⁵

87. In describing his use of supplemental evidence, Dr. Cole stated that he followed a professionally recognized methodology in analyzing and collecting supplemental evidence.⁵⁶ Despite that claim, nothing in his preliminary report, rebuttal report, or deposition offered anything to support the assertion that his collection of qualitative ‘supplemental’ data in this case was in any sense systematic or that it followed a clear protocol explicitly established in advance and adhered to carefully in its application.

88. In his rebuttal report, Dr. Cole asserts, for the first time, that he collected his supplemental evidence as part of a “mixed-methods approach,” in which qualitative data is “used to improve the quantitative study by providing deeper and fuller answers to the research question.”⁵⁷ But apart from a simple name-check, Dr. Cole does not actually describe the substance of his mixed-methodology approach—he does not explain the role qualitative data played in his analysis, nor does he provide a method for resolving conflicts between the different

⁵⁴ See *id.* at 45.

⁵⁵ Cole Report at 13; Cole Rebuttal Report at 4.

⁵⁶ Cole Tr. at 87.

⁵⁷ Cole Rebuttal Report at 5.

types of information. Each of those explanations is an essential part of a mixed-methodology research plan, and each part is needed to prevent mixed-methodology inquiries from devolving into a subjective, standardless inquiry.

VIII. Conclusion

89. Based on the above analysis, I conclude that Dr. Cole's conclusions regarding voting patterns in District elections are not reliable. None of Dr. Cole's methods—whether viewed individually or in combination—supports his conclusion that an empirical analysis of East Ramapo elections demonstrates that minorities are voting cohesively, that there is racially polarized voting, or that minority voters are unable to elect their candidates of choice.

90. Pursuant to 28 U.S.C. § 1746, I declare under penalty of perjury of the laws of the United States that the foregoing is true and correct to the best of my knowledge, information, and belief.

Date: March 28, 2018

A handwritten signature in black ink, appearing to be 'Joseph Cole', written in a cursive style.